# Supporting Information

## Production of Tagatose-Rich Syrup with Ion-Exchange Resin Purification

Nilofar Arabi,[1] Seyedamirreza Babaei,[1] Ji Qi,[1] Shubhangi Arvelli,[1] Shaghayegh Janbazialamdari,[1] Jikai Zhao[1]*

[1]Carl and Melinda Helwig Department of Biological and Agricultural Engineering, Kansas State University, Manhattan, KS 66506, USA

***Corresponding author**: Dr. Jikai Zhao (jikaizhao@ksu.edu)

## S1. Modeling and Optimization of Resin Dosage

This section provides the regression methodology, model formulation, and experimental dataset used to quantify the effects of resin dosage and contact time on syrup deionization performance. These details complement the summarized results presented in Section 3.6 of the main manuscript.

### S1.1 Regression Methodology

Experimental data from batch ion-exchange runs were normalized by syrup volume and expressed as cation ($D_c$) and anion ($D_a$) resin loadings (g resin $mL^{-1}$ syrup). Syrup volumes ranged from 10 to 300 mL, with contact times of 5-60 min per exchange stage and inlet conductivities near 2.3-2.4 mS $cm^{-1}$. Conductivity values and syrup volume were $\log_{10}$-transformed prior to regression. Model coefficients were estimated using ordinary least squares regression implemented in Python (NumPy, Python 3.11).[1] All predictor variables were normalized prior to fitting.

### S1.2 Polynomial Regression Models

#### S1.2.1 Cation-exchange regression

A second-order polynomial regression with interaction terms was used to relate the logarithm of conductivity after cation exchange to the logarithm of inlet conductivity, cation resin dosage ($D_c$), contact time ($t_c$), and syrup volume (V). The fitted relation was:

$$\log_{10}(cond_{cat}) = f(\log_{10}(cond_{in}), D_c, t_c, \log_{10}(V)) \quad (1)$$

The model achieved an $R^2$ of 0.87, capturing the nonlinear behavior associated with cation-exchange saturation.

#### S1.2.2 Anion-exchange regression

An analogous second-order polynomial regression was applied to the anion-exchange step using the logarithm of conductivity exiting the cation stage, anion resin dosage ($D_a$), contact time ($t_a$), and syrup volume to predict final syrup conductivity:

$$\log_{10}(cond_{out}) = f(\log_{10}(cond_{cat}), D_a, t_a, \log_{10}(V)) \quad (2)$$

The anion-exchange model produced an $R^2$ of 0.92, indicating strong predictive capability and nonlinear saturation behavior.

### S1.3 Experimental dataset

**Table S1.** Experimental conditions and measured responses for batch cation- and anion-exchange deionization experiments used for regression modeling.

| Run | Syrup (mL) | ($t_c$) (min) | ($t_a$) (min) | ($m_c$) (g) | ($m_a$) (g) | ($D_c$) (g $mL^{-1}$) | ($D_a$) (g $mL^{-1}$) | Cond in (µS $cm^{-1}$) | Cond after cat (µS $cm^{-1}$) | Cond out (µS $cm^{-1}$) | pH out |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 200 | 5 | 10 | 7.76 | 9.09 | 0.04 | 0.05 | 2559 | 2411 | 1443 | 7.36 |
| 2 | 300 | 15 | 25 | 14.13 | 1.97 | 0.05 | 0.01 | 2557 | 2204 | 1267 | 6.84 |
| 3 | 10 | 10 | 25 | 1.22 | 0.72 | 0.12 | 0.07 | 2449 | 1722 | 1978 | 8.36 |

| Run | Syrup (mL) | ($t_c$) (min) | ($t_a$) (min) | ($m_c$) (g) | ($m_a$) (g) | ($D_c$) (g $mL^{-1}$) | ($D_a$) (g $mL^{-1}$) | Cond in (µS $cm^{-1}$) | Cond after cat (µS $cm^{-1}$) | Cond out (µS $cm^{-1}$) | pH out |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 100 | 20 | 30 | 3.92 | 4.08 | 0.04 | 0.04 | 2506 | 2022 | 1765 | 8.04 |
| 5 | 200 | 25 | 5 | 24.07 | 7.02 | 0.12 | 0.04 | 2386 | 1512 | 843 | 7.19 |
| 6 | 300 | 10 | 25 | 39.11 | 20.02 | 0.13 | 0.07 | 2558 | 956 | 412 | 8.2 |
| 7 | 300 | 30 | 30 | 29.62 | 15.15 | 0.10 | 0.05 | 2564 | 1112 | 527 | 7.98 |
| 8 | 50 | 25 | 5 | 6.68 | 0.43 | 0.13 | 0.01 | 2527 | 1538 | 1432 | 7.11 |
| 9 | 50 | 20 | 15 | 1.36 | 1.35 | 0.03 | 0.03 | 2540 | 2270 | 1886 | 6.69 |
| 10 | 10 | 30 | 20 | 1.04 | 0.60 | 0.10 | 0.06 | 2307 | 1460 | 704 | 7.84 |
| 11 | 200 | 25 | 5 | 26.62 | 1.36 | 0.13 | 0.01 | 2347 | 1457 | 1312 | 6.9 |
| 12 | 300 | 10 | 15 | 18.10 | 22.59 | 0.06 | 0.08 | 2372 | 1882 | 1029 | 7.8 |
| 13 | 300 | 30 | 5 | 38.39 | 5.57 | 0.13 | 0.02 | 2377 | 1368 | 1180 | 6.92 |
| 14 | 300 | 30 | 30 | 50.13 | 7.94 | 0.17 | 0.03 | 2410 | 865 | 246 | 7.6 |
| 15 | 300 | 30 | 5 | 35.00 | 17.01 | 0.12 | 0.06 | 2334 | 1395 | 451 | 7.7 |
| 16 | 10 | 10 | 10 | 0.78 | 0.79 | 0.08 | 0.08 | 2372 | 1805 | 2111 | 8.29 |
| 17 | 25 | 5 | 5 | 1.02 | 1.36 | 0.04 | 0.05 | 2330 | 1429 | 373 | 7.08 |
| 18 | 300 | 30 | 20 | 36.80 | 2.30 | 0.12 | 0.01 | 2368 | 1552 | 630 | 7.2 |
| 19 | 50 | 30 | 5 | 6.18 | 3.29 | 0.12 | 0.07 | 2311 | 1254 | 388 | 7.39 |
| 20 | 100 | 5 | 15 | 5.11 | 6.89 | 0.05 | 0.07 | 2381 | 2103 | 1447 | 8.28 |

**Table S1** summarizes the experimental conditions and measured responses used for regression and model validation, including syrup volume, resin dosages, contact times, conductivities at each processing stage, and outlet pH.

## S2. Supplementary microbiological data

This section provides raw plate count data, dilution factors, and calculated microbial counts used to determine the microbiological quality of concentrated tagatose-galactose syrups at 65 and 75 °Brix. Duplicate plate counts were obtained for total aerobic mesophilic bacteria, yeast and mold, and selected foodborne pathogens following FDA Bacteriological Analytical Manual (BAM) procedures.[2] The data presented in Table S2 support the summarized microbial results reported in Section 3.5.5 of the main manuscript and allow transparent verification of CFU calculations.

**Table S2.** Raw microbiological enumeration data and dilution factors used for CFU calculations in concentrated tagatose-galactose syrups.

| Sample (°Brix) | Microorganism | Plate count (replicate 1) | Plate count (replicate 2) | Dilution factor | Calculated CFU/mL | Log CFU/mL |
|---|---|---|---|---|---|---|
| **65 °Brix** | Total aerobic mesophilic count (TAMC) | 25 | 25 | $10^0$ | 37.5 | 1.57 |
| Yeast and mold | 15 | 15 | $10^{-1}$ | 225 | 2.35 | |
| *Escherichia coli* | N.D. | N.D. | — | N.D. | — | |
| *Salmonella* spp. | N.D. | N.D. | — | N.D. | — | |
| *Staphylococcus aureus* | N.D. | N.D. | — | N.D. | — | |
| **75 °Brix** | Total aerobic mesophilic count (TAMC) | 26 | 30 | $10^0$ | 43 | 1.63 |
| Yeast and mold | 15 | 20 | $10^{-1}$ | 275 | 2.44 | |
| *Escherichia coli* | N.D. | N.D. | — | N.D. | — | |
| *Salmonella* spp. | N.D. | N.D. | — | N.D. | — | |
| *Staphylococcus aureus* | N.D. | N.D. | — | N.D. | — | |

**Notes:** N.D. , not detected at the limit of detection of the applied method. CFU $mL^{-1}$ values were calculated from observed plate counts and corresponding dilution factors according to standard plate count methods. Log CFU $mL^{-1}$ values represent $log_{10}$-transformed counts. Microbiological analyses were conducted after approximately two weeks of refrigerated storage (4 °C) to evaluate microbial stability of the concentrated syrups. All measurements were performed in duplicate.

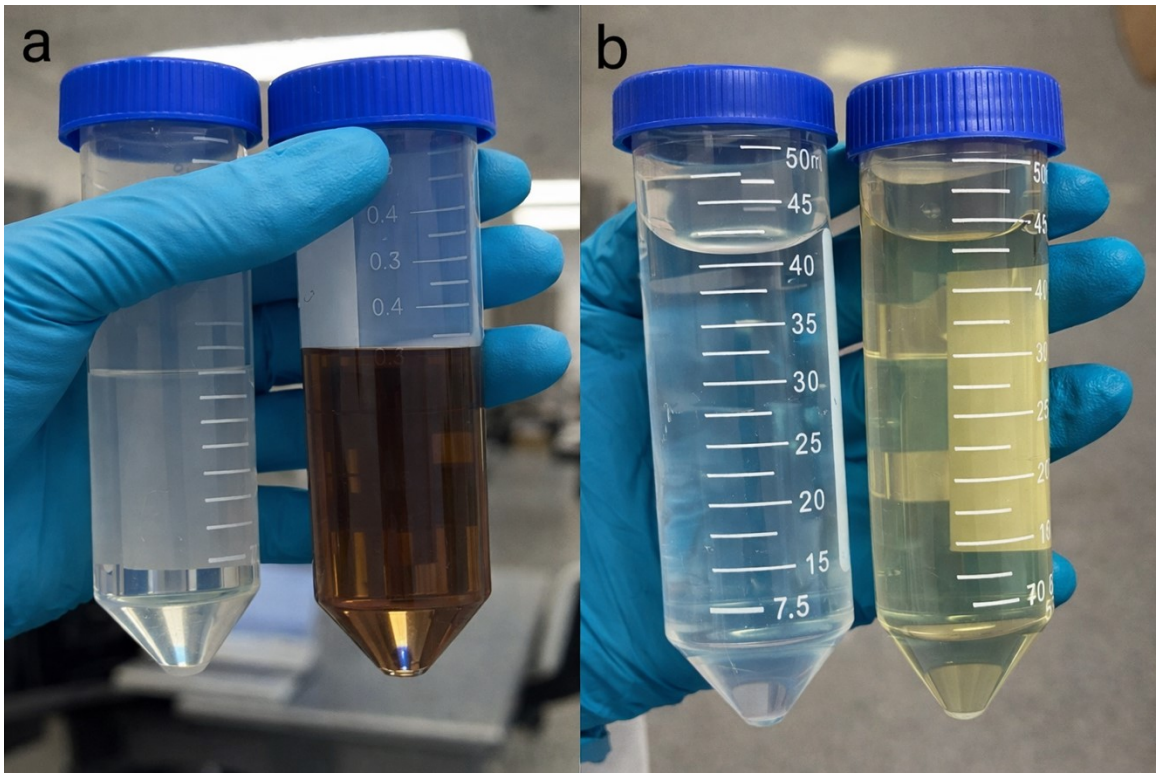


**Figure S1.** Visual comparison of syrup color after neutralization using different treatments: (a) $H_2SO_4$-treated syrup and (b) $CO_2$-treated syrup.

## References

1 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. VanderPlas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, J. Mach. Learn. Res., 2011, 12, 2825–2830.

2 Bacteriological Analytical Manual (BAM) | FDA, https://www.fda.gov/food/laboratory-methods-food/bacteriological-analytical-manual-bam.